How can I extract the Root Mean Squared Error (RMSE) from the lm() function in R?

Authored by

stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I extract the Root Mean Squared Error (RMSE) from the lm() function in R?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=154178

The lm() function in R is used to fit linear regression models. It calculates the predicted values for a given dataset and compares them to the actual values, resulting in a residual error. The Root Mean Squared Error (RMSE) is a commonly used measure of this error, which represents the standard deviation of the residuals. To extract the RMSE from the lm() function in R, the summary() function can be used to obtain a summary of the model, including the RMSE value. This value can then be accessed and used for further analysis or comparison with other models. Overall, the lm() function in R provides a simple and efficient way to calculate and extract the RMSE for linear regression models.

Extract RMSE from lm() Function in R

You can use the following syntax to extract the root mean square error (RMSE) from the function in R:

sqrt(mean(model$residuals^2))

The following example shows how to use this syntax in practice.

Related:

Example: Extract RMSE from lm() in R

Suppose we fit the following multiple linear regression model in R:

#create data frame
df <- data.frame(rating=c(67, 75, 79, 85, 90, 96, 97),
points=c(8, 12, 16, 15, 22, 28, 24),

```
assists=c(4, 6, 6, 5, 3, 8, 7),
rebounds=c(1, 4, 3, 3, 2, 6, 7))

#fit multiple linear regression model
model <- lm(rating ~ points + assists + rebounds, data=df)
```

We can use the summary() function to view the entire summary of the regression model:

```
#view model summary
summary(model)

Call:
lm(formula = rating ~ points + assists + rebounds, data = df)

Residuals:
1 2 3 4 5 6 7
-1.5902 -1.7181 0.2413 4.8597 -1.0201 -0.6082 -0.1644

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 66.4355 6.6932 9.926 0.00218 **
points 1.2152 0.2788 4.359 0.02232 *
```

assists -2.5968 1.6263 -1.597 0.20860

rebounds 2.8202 1.6118 1.750 0.17847

Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Residual standard error: 3.193 on 3 degrees of freedom

Multiple R-squared: 0.9589, Adjusted R-squared: 0.9179

F-statistic: 23.35 on 3 and 3 DF, p-value: 0.01396

To only extract the root mean square error (RMSE) of the model, we can use the following syntax:

#extract RMSE of regression model

sqrt(mean(model$residuals^2))

2.090564

The RMSE of the model is 2.090564.

This represents the average distance between the predicted values from the model and the actual values in the dataset.

Note that the lower the RMSE, the better a given model is able to "fit" a dataset.

When comparing several different regression models, the model with the lowest RMSE is said to be the one that "fits" the dataset the best.